WHAT IS CLAIMED IS:

1. An isolated nucleic acid comprising a sequence encoding a fibrocystin polypeptide.

2. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide is encoded by SEQ ID NO:1.

3. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide is encoded by SEQ ID NO:3.

4. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide is encoded by SEQ ID NO:4.

5. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises the amino acid sequence of SEQ ID NO:2.

6. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises the amino acid sequence of SEQ ID NO:6.

7. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises the amino acid sequence of SEQ ID NO:7.

8. The isolated nucleic acid of claim 1, wherein said sequence comprises a sequence variant associated with autosomal recessive polycystic kidney disease.

9. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises an amino acid sequence variant at a position selected from the group consisting of: position 17, position 36, position 222, position 739, position 757, position 805, position 1249, position 1389, position 1407, position 1664, position 1741, position 1833, position 1838, position 1867, position 1917, position 1942, position 1995, position 2331, position 2688, position 2869, position 2957, position 3018, position 3177, position 3346, position 3468, position 3502, position 3529, position 3553, and position 3622 of SEQ ID NO:2.

10. The isolated nucleic acid of claim 9, wherein said amino acid sequence variant is selected from the group consisting of: Val at position 17, Met at position 36, Val at position 222, Leu at position 739, Leu at position 757, Leu at position 805, Trp at position 1249, Thr at position 1389, Arg at position 1407, Phe at position 1664, Met at position 1741, Leu at position 1833, Cys at position 1838, Asn at position 1867, Arg at position 1917, Gly at position 1942, Gly at position 1995, Lys at position 2331, Phe at position 2688, Lys at position 2869, Thr at position 2957, Phe at position 3018, Thr at position 3177, Arg at position 3346, Val at position 3468, Val at position 3502, Gln at position 3529, Thr at position 3553, and Tyr at position 3622.

11. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises an amino acid sequence variant at a position selected from the group consisting of: position 25, position 752, position 760, position 830, position 852, position 1262, position 1709, position 1870, position 2938, position 3139, position 3505, position 3899, position 3960, and position 4048 of SEQ ID NO:2.

12. The isolated nucleic acid of claim 11, wherein said amino acid sequence variant is selected from the group consisting of: Val at position 25, Met at position 752, Cys at position 760, Ser at position 830, Arg at position 852, Val at position 1262, Phe at position 1709, Val at position 1870, Met at position 2938, Tyr at position 3139, Arg at position 3505, Arg at position 3899, Ile at position 3960, and Arg at position 4048.

13. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises the amino acids from position 1 to 3299 of SEQ ID NO:2.

14. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises the amino acids from position 1 to 2578 of SEQ ID NO:2.

15. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide comprises the amino acids from position 1 to 3779 of SEQ ID NO:2.

16. The isolated nucleic acid of claim 1, wherein said sequence comprises a nucleotide sequence variant with respect to SEQ ID NO:1, SEQ ID NO:214, SEQ ID NO:216, or SEQ ID NO:217.

17. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 is at a position selected from the group consisting of: position 50, position 107, position 657, position 664, position 2216, position 2269, position 2414, position 3747, position 3761, position 4165, position 4220, position 4991, position 5221, position 5498, position 5513, position 5600, position 5750, position 5825, position 5984, position 6992, position 8011, position 8063, position 8606, position 8870, position 9053, position 9530, position 10036, position 10174, position 10402, position 10505, position 10585, position 10658, position 10865, and position 11612 of SEQ ID NO:1.

18. The isolated nucleic acid of claim 17, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 is selected from the group consisting of: T at position 50, T at position 107, T at position 657, G at position 664, T at position 2216, C at position 2269, T at position 2414, G at position 3747, G at position 3761, A at position 4165, G at position 4220, T at position 4991, A at position 5221, T at position 5498, G at position 5513, A at position 5600, G at position 5750, G at position 5825, G at position 5984, A at position 6992, T at position 8011, T at position 8063, A or T at position 8606, C at position 8870, T at position 9053, C at position 9530, C at position 10036, T at position 10174, G at position 10402, T at position 10505, C at position 10585, C at position 10658, A at position 10865, and A at position 11612.

19. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises an A inserted at position 5895 or 5896.

20. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotides at positions 1624, 1625, 1626, and 1627.

21. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 10637.

22. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 9689.

23. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 3762.

24. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 1529.

25. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 383.

26. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 6383.

27. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 10856.

28. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 comprises a deletion of the nucleotide at position 10364.

29. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:214 is at position -2 relative to the splice acceptor site of intron 28.

30. The isolated nucleic acid of claim 29, wherein said nucleotide sequence variant is a C at position -2 relative to the splice acceptor site of intron 28.

31. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:216 is at position -9 relative to the splice acceptor site of intron 33.

32. The isolated nucleic acid of claim 31, wherein said nucleotide sequence variant is a G at position -9 relative to the splice acceptor site of intron 33.

33. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:217 is at position +4 relative to the splice donor site of intron 43.

34. The isolated nucleic acid of claim 33, wherein said nucleotide sequence variant is a T at position +4 relative to the splice donor site of intron 43.

35. The isolated nucleic acid of claim 16, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 is at a position selected from the group consisting of: position 73, position 214, position 234, position 1185, position 1587, position 2046, position 2196, position 2255, position 2278, position 2489, position 2554, position 2853, position 3537, position 3756, position 3785, position 4920, position 5125, position 5608, position 7587, position 7764, position 8813, position 9237, position 9415, position 10515, position 10521, position 11340, position 11196, position 11878, and position 12143 of SEQ ID NO:1.

36. The isolated nucleic acid of claim 35, wherein said nucleotide sequence variant with respect to SEQ ID NO:1 is selected from the group consisting of: A at position 73, T at position 214, T at position 234, C at position 1185, C at position 1587, C at position 2046, T at position 2196, T at position 2255, T at position 2278, G at position 2489, C at position 2554, T at position 2853, C at position 3537, C at position 3756, T at position 3785, G at position 4920, T at position 5125, G at position 5608, A at position 7587, G at position 7764, T at position 8813, A at position 9237, T at position 9415, T at position 10515, T at position 10521, C at position 11340, G at position 11196, A at position 11878, and G at position 12143.

37. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide is encoded by nucleotides 276 to 10174 of SEQ ID NO:1.

38. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide is encoded by nucleotides 276 to 8011 of SEQ ID NO:1.

39. The isolated nucleic acid of claim 1, wherein said fibrocystin polypeptide is encoded by nucleotides 276 to 11612 of SEQ ID NO:1.

40. The isolated nucleic acid of claim 1, said nucleic acid comprising nucleotides 1 to 192 of SEQ ID NO:1.

41. The isolated nucleic acid of claim 1, said nucleic acid comprising nucleotides 193 to 328 of SEQ ID NO:1.

42. The isolated nucleic acid of claim 1, said nucleic acid comprising nucleotides 329 to 406 of SEQ ID NO:1.

43. The isolated nucleic acid of claim 1, wherein said nucleotide sequence comprises a sequence variant with respect to SEQ ID NO:5, SEQ ID NO:209, SEQ ID NO:210, SEQ ID NO:211, SEQ ID NO:212, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:218, or SEQ ID NO:219.

44. The isolated nucleic acid of claim 43, wherein said nucleotide sequence variant is at a position selected from the group consisting of: position -47 relative to the splice acceptor site of SEQ ID NO:5, the position just 5' to the splice donor site of SEQ ID NO:209, position +19 relative to the splice donor site of SEQ ID NO:210, position +23 relative to the splice donor site of SEQ ID NO:211, position +13 relative to the splice donor site of SEQ ID NO:212, position +50 relative to the splice donor site of SEQ ID NO:213, position +53 relative to the splice donor site of SEQ ID NO:213, positions +42 through +45 relative to the splice donor site of SEQ ID NO:215, position -32 relative to the splice acceptor site of SEQ ID NO:218, and position +9 relative to the splice donor site of SEQ ID NO:219.

45. The isolated nucleic acid of claim 44, wherein said nucleotide sequence variant is a T at position -47 relative to the splice acceptor site of SEQ ID NO:5, an A inserted just 5' to the splice donor site of SEQ ID NO:209, a C at position +19 relative to the splice

donor site of SEQ ID NO:210, a T at position +23 relative to the splice donor site of SEQ ID NO:211, a G at position +13 relative to the splice donor site of SEQ ID NO:212, a T at position +50 relative to the splice donor site of SEQ ID NO:213, a G at position +53 relative to the splice donor site of SEQ ID NO:213, deletion of the nucleotides at positions +42 through +45 relative to the splice donor site of SEQ ID NO:215, a G at position -32 relative to the splice acceptor site of SEQ ID NO:218, or a G at position +9 relative to the splice donor site of SEQ ID NO:219.

46. An isolated nucleic acid encoding a fibrocystin polypeptide, wherein said nucleic acid comprises at least 300 contiguous nucleotides of SEQ ID NO:1 or a sequence variant thereof.

47. A vector comprising the isolated nucleic acid of claim 46.

48. Host cells comprising the vector of claim 47.

49. An isolated nucleic acid 10 to 1650 nucleotides in length, said nucleic acid comprising a sequence, said sequence comprising one or more nucleotide sequence variants relative to the sequence of SEQ ID NO:1, and wherein said sequence is at least 80% identical over its length to the corresponding sequence in SEQ ID NO:1.

50. The isolated nucleic acid of claim 49, wherein said one or more nucleotide sequence variants are at position 50, 107, 383, 657, 664, 1529, 1624, 1625, 1626, 1627, 2216, 2269, 2414, 3747, 3761, 3762, 4165, 4220, 4991, 5221, 5498, 5513, 5600, 5750, 5825, 5895, 5896, 5984, 6383, 6992, 8011, 8063, 8606, 8870, 9053, 9530, 9689, 10036, 10174, 10364, 10402, 10505, 10585, 10637, 10658, 10856, 10865, or 11612 of SEQ ID NO:1.

51. The isolated nucleic acid of claim 49, wherein said one or more nucleotide sequence variants are at position 73, 214, 234, 1185, 1587, 2046, 2196, 2255, 2278, 2489, 2554, 2853, 3537, 3756, 3785, 4920, 5125, 5608, 7587, 7764, 8813, 9237, 9415, 10515, 10521, 11340, 11196, 11878, or 12143 of SEQ ID NO:1.

52. A plurality of oligonucleotide primer pairs, wherein each primer is 10 to 50 nucleotides in length, and wherein each said primer pair, in the presence of mammalian genomic DNA and under polymerase chain reaction conditions, produces a nucleic acid product corresponding to a region of an ARPKD nucleic acid molecule, wherein said product is 30 to 1650 nucleotides in length.

53. The plurality of primer pairs of claim 52, wherein said nucleic acid product comprises a nucleotide sequence variant relative to SEQ ID NO:1.

54. The plurality of primer pairs of claim 52, wherein said plurality comprises at least three primer pairs.

55. The plurality of primer pairs of claim 52, wherein said plurality comprises at least thirteen primer pairs.

56. The plurality of primer pairs of claim 52, wherein said plurality comprises at least sixteen primer pairs.

57. The plurality of primer pairs of claim 52, wherein said plurality comprises at least twenty-three primer pairs.

58. A composition comprising a first oligonucleotide primer and a second oligonucleotide primer, wherein said first oligonucleotide primer and said second oligonucleotide primer are each 10 to 50 nucleotides in length, and wherein said first and second primers, in the presence of mammalian genomic DNA and under polymerase chain reaction conditions, produce a nucleic acid product corresponding to a region of an ARPKD nucleic acid molecule, wherein said product is 30 to 1650 nucleotides in length.

59. The composition of claim 58, wherein said nucleic acid product comprises a nucleotide sequence variant relative to SEQ ID NO:1.

60. An isolated nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or its complement.

61. An antibody having specific binding affinity for a fibrocystin polypeptide.

62. A method for determining the susceptibility of a subject to autosomal recessive polycystic kidney disease, said method comprising providing a nucleic acid sample from said subject, and determining whether said nucleic acid sample contains one or more nucleotide sequence variants within the *PKHD1* gene of said subject relative to a wild-type *PKHD1* gene, wherein the presence of said one or more sequence variants is associated with increased susceptibility of said subject to autosomal recessive polycystic kidney disease.

63. The method of claim 62, wherein said nucleic acid sample is genomic DNA.

64. The method of claim 62, wherein said determining step is performed by denaturing high performance liquid chromatography.

65. The method of claim 62, further comprising identifying said one or more sequence variants by DNA sequencing.

66. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion of the nucleotides at positions 1624, 1625, 1626, and 1627, and an A at position 6992 of SEQ ID NO:1.

67. The method of claim 62, wherein said one or more nucleotide sequence variants are a G at position 664 and a T at position 10174 of SEQ ID NO:1.

68. The method of claim 62, wherein said one or more nucleotide sequence variants are a G at position 4220 and an A inserted at position 5896 of SEQ ID NO:1.

69. The method of claim 62, wherein said one or more nucleotide sequence variants are a T at position 8011 and a C at position 10658 of SEQ ID NO:1.

70. The method of claim 62, wherein said one or more nucleotide sequence variants are a G at position 5984 and an A at position 11612 of SEQ ID NO:1.

71. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 10637, and a C at position 8870 of SEQ ID NO:1.

72. The method of claim 62, wherein said one or more nucleotide sequence variants are a T at position 4991 and a T at position 9053 of SEQ ID NO:1.

73. The method of claim 62, wherein said one or more nucleotide sequence variants are a G at position 3747 and a G at position 5750 of SEQ ID NO:1.

74. The method of claim 62, wherein said one nucleotide sequence variant is an A at position 5221 of SEQ ID NO:1.

75. The method of claim 62, wherein said one nucleotide sequence variant is a T at position 107 of SEQ ID NO:1.

76. The method of claim 62, wherein said one nucleotide sequence variant is a deletion at position 9689 of SEQ ID NO:1.

77. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 9689 and a G at position 3761 in combination with a deletion at position 3762 of SEQ ID NO:1.

78. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 9689 and an A at position 10865 of SEQ ID NO:1.

79. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 9689 and a T at position 50 of SEQ ID NO:1.

80. The method of claim 62, wherein said one or more nucleotide sequence variants are an A inserted at position 5895, a T at position 8063, and a G at position 10402 of SEQ ID NO:1.

81. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 1529, a T at position 657, and an A at position 8606 of SEQ ID NO:1.

82. The method of claim 62, wherein said one or more nucleotide sequence variants are a G at position 664 and a G at position 3761 in combination with a deletion at position 3762 of SEQ ID NO:1.

83. The method of claim 62, wherein said one or more nucleotide sequence variants are an insertion at position 5895 and a C at position 10036 of SEQ ID NO:1.

84. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 383 and a G at position 5513 of SEQ ID NO:1.

85. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 6383 and a G at position 664 of SEQ ID NO:1.

86. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 383 and a G at position 664 of SEQ ID NO:1.

87. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 10856 of SEQ ID NO:1 and a G at position -9 relative to the splice acceptor site of SEQ ID NO:216.

88. The method of claim 62, wherein said one or more nucleotide sequence variants are a T at position 10505 and an A at position 8606 of SEQ ID NO:1, and a C at position -2 relative to the splice acceptor site of SEQ ID NO:214.

89. The method of claim 62, wherein said one or more nucleotide sequence variants are a T at position 107 of SEQ ID NO:1 and a T at position +4 relative to the splice donor site of SEQ ID NO:217.

90. The method of claim 62, wherein said one or more nucleotide sequence variants are a G at position 5825, a T at position 8606, and a T at position 2216 of SEQ ID NO:1.

91. The method of claim 62, wherein said one or more nucleotide sequence variants are a T at position 107, a T at position 2414, and a C at position 9530 of SEQ ID NO:1.

92. The method of claim 62, wherein said one or more nucleotide sequence variants are a C at position 2269 and a C at position 9530 of SEQ ID NO:1.

93. The method of claim 62, wherein said one nucleotide sequence variant is a deletion at position 1529 of SEQ ID NO:1.

94. The method of claim 62, wherein said one nucleotide sequence variant is an A inserted at position 5895 of SEQ ID NO:1.

95. The method of claim 62, wherein said one nucleotide sequence variant is an A at position 5600 of SEQ ID NO:1.

96. The method of claim 62, wherein said one nucleotide sequence variant is a C at position 10585 of SEQ ID NO:1.

97. The method of claim 62, wherein said one nucleotide sequence variant is an A at position 4165 of SEQ ID NO:1.

98. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 9689 and an A at position 8606 of SEQ ID NO:1.

99. The method of claim 62, wherein said one or more nucleotide sequence variants are a deletion at position 10364 and a G at position 10402 of SEQ ID NO:1.

100. The method of claim 62, wherein said one or more nucleotide sequence variants are an A at position 5221 and a T at position 5498 of SEQ ID NO:1.

101. The method of claim 62, wherein said one or more nucleotide sequence variants are an A at position 8606 and a C at position 8870 of SEQ ID NO:1.

102. The method of any of claims 62 to 101, wherein said one or more nucleotide sequence variants are on separate alleles.

103. A method for diagnosing autosomal recessive polycystic kidney disease in a subject, said method comprising providing a nucleic acid sample from said subject, and determining whether said nucleic acid sample contains one or more disease-associated

sequence variants within the *PKHD1* gene of said subject compared to a wild-type *PKHD1* gene, wherein the presence of said one or more disease-associated sequence variants is diagnostic of autosomal recessive polycystic kidney disease.

104. An article of manufacture comprising a substrate, wherein said substrate comprises a population of isolated nucleic acid molecules, wherein each said nucleic acid molecule is 10 to 1000 nucleotides in length, wherein each said nucleic acid molecule comprises a different nucleotide sequence variant relative to the sequence of SEQ ID NO:1, and wherein said nucleic acid molecule is at least 80% identical over its length to the corresponding sequence in SEQ ID NO:1.